

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-229740

(43)Date of publication of application : 16.08.2002

(51)Int.CI.

G06F 3/06

G06F 12/16

G06F 13/00

(21)Application number : 2001-023974

(71)Applicant : TOSHIBA CORP

(22)Date of filing : 31.01.2001

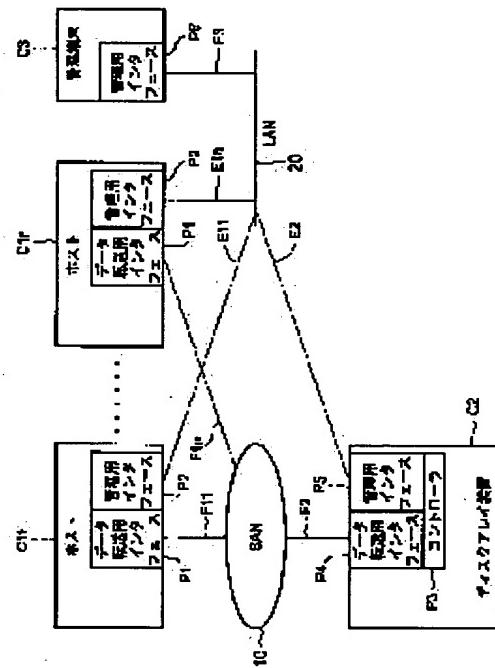
(72)Inventor : YAMASHITA AKIRA

(54) STORAGE SYSTEM WITH HIGH AVAILABILITY CONTROL INTERFACE, HOST COMPUTER USED THEREFOR, AND COMPUTER STORED WITH DRIVER PROGRAM EXECUTED BY COMPUTER

(57)Abstract:

PROBLEM TO BE SOLVED: To reduce the number of components necessary for multiplexing of data transfer path by sharing a control path for a data transfer path.

SOLUTION: A duplex driver in a host computer C11, controls in data transferring between the host computer C11 and a disk array device C2, by switching a path F led to a data transfer interface P4 of a disk array device C2 via a cable F11, SAN 10 and a cable F2 from a data transfer interface P1 of the host computer C11 and a path E from a control interface P2 of the host computer C11 led to a control interface P5 of the disk array device C2 via a cable E11, LAN 20 and a cable E2, as a data transfer path.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's

[decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

管理用インターフェースを用いた高可用性ストレージシステム、同システムに用いられるホ...
特開2002-229740

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開2002-229740

(P2002-229740A)

(43)公開日 平成14年8月16日(2002.8.16)

(51)Int.Cl. ⁷	識別記号	F I	テマコト [*] (参考)
G 0 6 F 3/06	3 0 1	G 0 6 F 3/06	3 0 1 B 5 B 0 1 8
12/16	3 2 0	12/16	3 2 0 L 5 B 0 6 5
13/00	3 0 1	13/00	3 0 1 P 5 B 0 8 3

審査請求 未請求 請求項の数10 O.L (全 11 頁)

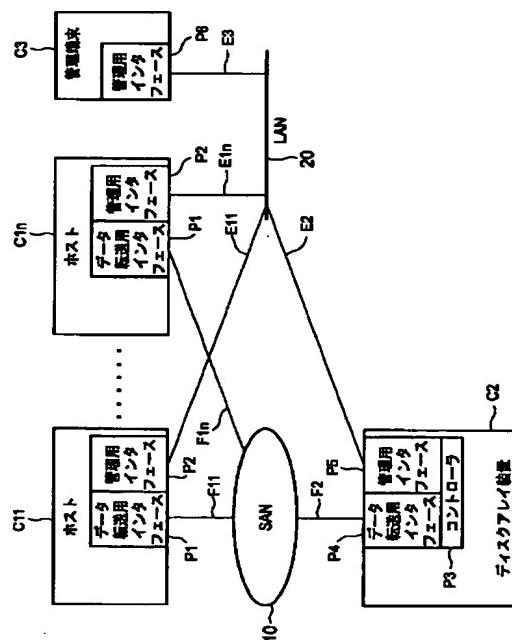
(21)出願番号	特願2001-23974(P2001-23974)	(71)出願人	000003078 株式会社東芝 東京都港区芝浦一丁目1番1号
(22)出願日	平成13年1月31日(2001.1.31)	(72)発明者	山下亮 東京都府中市東芝町1番地 株式会社東芝 府中事業所内
		(74)代理人	100058479 弁理士 鈴江 武彦 (外6名) Fターム(参考) 5B018 GA06 HA35 HA40 MA14 5B065 BA01 CA05 CA30 5B083 AA08 BB03 CC04 CD11 EE08

(54)【発明の名称】 管理用インターフェースを用いた高可用性ストレージシステム、同システムに用いられるホストコンピュータ、及び同コンピュータにより実行されるドライバプログラムを記憶したコンピュータ

(57)【要約】

【課題】 管理用バスをデータ転送用バスに共用することにより、データ転送用バスの多重化に必要な構成要素の削減を図る。

【解決手段】 ホストC11内の二重化ドライバは、ホストC11とディスクアレイ装置C2との間のデータ転送で、そのデータ転送用のバスとして、ホストC11のデータ転送用インターフェースP1からケーブルF11、SAN10及びケーブルF2を経由してディスクアレイ装置C2のデータ転送用インターフェースP4に至るバスFと、ホストC11の管理用インターフェースP2からケーブルE11、LAN20及びケーブルE2を経由してディスクアレイ装置C2の管理用インターフェースP5に至るバスEとを切り替え制御する。



【特許請求の範囲】

【請求項1】 データ転送のための第1のインターフェースと、状態監視を含むシステム管理のための第2のインターフェースとを有する少なくとも1つのホストコンピュータと、

データ転送のための第3のインターフェースと、状態監視を含むシステム管理のための第4のインターフェースとを有する少なくとも1つのストレージ装置と、

前記ホストコンピュータの前記第1のインターフェースと前記ストレージ装置の前記第3のインターフェースとを接続するデータ転送用接続インターフェースと、

前記ホストコンピュータの前記第2のインターフェースと前記ストレージ装置の前記第4のインターフェースとを接続する管理用接続インターフェースとを具備し、

前記ホストコンピュータは、当該コンピュータと前記ストレージ装置との間のデータ転送で、そのデータ転送用のバスとして、当該コンピュータの前記第1のインターフェースから前記データ転送用接続インターフェースを経由して前記ストレージ装置の前記第3のインターフェースに至る第1のバスと、当該コンピュータの前記第2のインターフェースから前記管理用接続インターフェースを経由して前記ストレージ装置の前記第4のインターフェースに至る第2のバスとを切り替え制御する二重化ドライバを有することを特徴とする高可用性ストレージシステム。

【請求項2】 前記ホストコンピュータは、前記第2のインターフェース及び前記管理用接続インターフェースを利用しての管理データ転送と、前記管理用接続インターフェースを含む前記第2のバスを利用しての前記ストレージ装置との間のデータ転送とを制御する管理用ドライバを有することを特徴とする請求項1記載の高可用性ストレージシステム。

【請求項3】 前記ストレージ装置は、前記第4のインターフェース及び前記管理用接続インターフェースを利用しての管理データ転送と、前記管理用接続インターフェースを含む前記第2のバスを利用しての前記ホストコンピュータとの間のデータ転送とを制御する管理用ドライバを有することを特徴とする請求項1記載の高可用性ストレージシステム。

【請求項4】 前記ホストコンピュータの管理用ドライバは、前記第2のバスを利用してのデータ転送では、前記ストレージ装置に転送すべき前記データ転送用接続インターフェースでのデータ転送に適用されるデータ転送用プロトコルのディスクアクセスコマンド及びディスクアクセスデータを、前記管理用接続インターフェースを介して前記ストレージ装置に転送するための特定データ形式に変換すると共に、前記第2のインターフェースにより前記データ転送用接続インターフェースから受信された前記特定データ形式のデータを前記データ転送用プロトコルのデータに変換することを特徴とする請求項2記載の高可用性ストレージシステム。

【請求項5】 前記ストレージ装置の管理用ドライバは、前記第2のバスを利用してのデータ転送では、前記ホストコンピュータに転送すべき前記データ転送用接続インターフェースでのデータ転送に適用されるデータ転送用プロトコルのディスクアクセスコマンド及びディスク

アクセスデータを、前記管理用接続インターフェースを介して前記ホストコンピュータに転送するための特定データ形式に変換すると共に、前記第4のインターフェースにより前記データ転送用接続インターフェースから受信された前記特定データ形式のデータを前記データ転送用プロトコルのデータに変換することを特徴とする請求項3記載の高可用性ストレージシステム。

【請求項6】 データ転送のための第1のインターフェースと、状態監視を含むシステム管理のための第2のインターフェースとを有し、データ転送のための第3のインターフェースと、状態監視を含むシステム管理のための第4のインターフェースとを有するストレージ装置と、前記第1のインターフェース、データ転送用接続インターフェース、及び前記第3のインターフェースにより接続されると

共に、前記第2のインターフェース、管理用接続インターフェース、及び前記第4のインターフェースにより接続される、高可用性ストレージシステムに用いられるホストコンピュータであって、前記コンピュータと前記ストレージ装置との間のデータ転送で、そのデータ転送用のバス

として、当該コンピュータの前記第1のインターフェースから前記データ転送用接続インターフェースを経由して前記ストレージ装置の前記第3のインターフェースに至る第1のバスと、当該コンピュータの前記第2のインターフェースから前記管理用接続インターフェースを経由して前記ストレージ装置の前記第4のインターフェースに至る第2のバスとを切り替え制御する二重化ドライバを具備することを特徴とする高可用性ストレージシステムに用いられるホストコンピュータ。

【請求項7】 前記第2のインターフェース及び前記管理用接続インターフェースを利用しての管理データ転送と、前記管理用接続インターフェースを含む前記第2のバスを利用しての前記ストレージ装置との間のデータ転送とを制御する管理用ドライバを更に具備することを特徴とする請求項6記載の高可用性ストレージシステムに用いられるホストコンピュータ。

【請求項8】 データ転送のための第1のインターフェースと、状態監視を含むシステム管理のための第2のインターフェースとを有し、データ転送のための第3のインターフェースと、状態監視を含むシステム管理のための第4のインターフェースとを有するストレージ装置と、前記第1のインターフェース、データ転送用接続インターフェース、及び前記第3のインターフェースにより接続されると共に、前記第2のインターフェース、管理用接続インターフェース、及び前記第4のインターフェースにより接続される、高可用性ストレージシステムに用いられるホストコ

ンピュータにより実行されるドライバプログラムを記憶したコンピュータ読み取り可能な記憶媒体であって、前記コンピュータを、

前記コンピュータの前記第1のインターフェースから前記データ転送用接続インターフェースを経由して前記ストレージ装置の前記第3のインターフェースに至る第1のバスと、前記コンピュータの前記第2のインターフェースから前記管理用接続インターフェースを経由して前記ストレージ装置の前記第4のインターフェースに至る第2のバスとを監視する監視手段と、

前記コンピュータと前記ストレージ装置との間のデータ転送で、そのデータ転送用のバスとして、前記第1のバスと前記第2のバスとを前記監視ステップの監視結果に基づいて切り替え制御する切り替え制御手段として機能させるためのドライバプログラムを記憶したコンピュータ読み取り可能な記憶媒体。

【請求項9】 前記コンピュータを、

前記第2のインターフェース及び前記管理用接続インターフェースを利用しての管理データ転送と、前記管理用接続インターフェースを含む前記第2のバスを利用しての前記ストレージ装置との間のデータ転送とを制御する転送制御手段として機能させるための請求項8記載のドライバプログラムを記憶したコンピュータ読み取り可能な記憶媒体。

【請求項10】 データ転送のための第1のインターフェースと、状態監視を含むシステム管理のための第2のインターフェースとを有し、データ転送のための第3のインターフェースと、状態監視を含むシステム管理のための第4のインターフェースとを有するストレージ装置と、前記第1のインターフェース、データ転送用接続インターフェース、及び前記第3のインターフェースにより接続されると共に、前記第2のインターフェース、管理用接続インターフェース、及び前記第4のインターフェースにより接続される、高可用性ストレージシステムに用いられるホストコンピュータにおける二重化バス制御方法であって、前記コンピュータの前記第1のインターフェースから前記データ転送用接続インターフェースを経由して前記ストレージ装置の前記第3のインターフェースに至る第1のバスと、前記コンピュータの前記第2のインターフェースから前記管理用接続インターフェースを経由して前記ストレージ装置の前記第4のインターフェースに至る第2のバスとを監視するステップと、

前記コンピュータと前記ストレージ装置との間のデータ転送で、そのデータ転送用のバスとして、前記第1のバスと前記第2のバスとを前記監視ステップの監視結果に基づいて切り替え制御するステップとを具備することを特徴とする二重化バス制御方法

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、ストレージ装置ア

クセスのためのデータ転送用のバスと、状態監視を含むシステム管理用のバスとを独立して有する高可用性ストレージシステム、同システムに用いられるホストコンピュータ、及び同コンピュータにより実行されるドライバプログラムを記憶したコンピュータ読み取り可能な記憶媒体に関する。

【0002】

【従来の技術】 一般に、ストレージ装置、例えば複数のディスク装置を備えたディスクアレイ装置は、ホストコンピュータ（以下、ホストと略称する）からのデータアクセスに用いられるインターフェース（データ転送用インターフェース）として、SCSI (Small Computer System Interface)、或いはファイバチャネルなどのインターフェースを少なくとも1つ実装している。

【0003】 ディスクアレイ装置を用いて構成されるシステム（ストレージシステム）には、高可用性を実現するために、ホストとディスクアレイ装置との間を複数のバス、例えば2つのバス（データ転送用バス）で接続しているものがある。このようなシステムは高可用性スト

レージシステムと呼ばれ、ホストとディスクアレイ装置のそれぞれに、バス数分のデータ転送用インターフェースが実装される。この高可用性ストレージシステムでは、稼動中のバス（実行バス）にケーブルの断線などの何らかの障害が発生した場合に、別のバス（待機バス）を使用して運用を継続させることができるように、複数のバスの運用をホストにて制御することによって、システムの可用性を向上させている。

【0004】

【発明が解決しようとする課題】 上記したように、従来のディスクアレイ装置を備えた高可用性ストレージシステムでは、高可用性を実現するために、ホストとディスクアレイ装置にそれぞれデータ転送用インターフェースを複数、例えば2つ実装し、ホストとディスクアレイ装置との間を2つのデータ転送用バスで接続して二重化バス構成とすることで、データ転送用バスの障害に対応できるようしている。

【0005】 このようなストレージシステムでは、2つのデータ転送用バスの一方に障害が発生しても、もう一方のバスを使用することで、確かにシステムの可用性を向上することができる。しかし、ホストとディスクアレイ装置にそれぞれデータ転送用インターフェースを2つ実装して二重化バス構成とすることは、ハードウェア構成要素数の増加を招くという問題がある。

【0006】 一方、高可用性ストレージシステムに用いられるディスクアレイ装置には、電源、冷却用ファンなどの構成要素の状態や、データのアクセス状況などの状態を外部の管理端末或いは管理用ホストに通知するための、例えばCSMA/CD (Carrier Sense Multiple Access with Collision Detection) 方式を適用する管理

用インターフェースを実装しているものがある。特に近時

は、システムの管理のために、ディスクアレイ装置だけでなく、ホストを含めて、全てのシステム構成機器の状態監視が行えるように、各システム構成機器に管理用インターフェースを実装して LAN (ローカルエリアネットワーク) と接続することにより管理用バスを設定し、各システム構成機器の状態監視を当該管理用バスを介して 1 つの管理端末或いは 1 つの管理用ホスト (ネットワーク接続機器) にて集中して行える高可用性ストレージシステムが実用化されている。

【0007】本発明は上記事情を考慮してなされたものでその目的は、ストレージ装置を含むシステム構成機器の状態監視等に用いられる管理用バスをホストとストレージ装置との間のデータ転送用バスに共用することにより、データ転送用バスの多重化に必要な構成要素を削減でき、構成要素を削減しない場合には多重化数を 1 つ増加してより一層可用性が向上できる高可用性ストレージシステム、同システムに用いられるホストコンピュータ、及び同コンピュータにより実行されるドライバプログラムを記憶したコンピュータ読み取り可能な記憶媒体を提供することにある。

【0008】

【課題を解決するための手段】本発明に係る高可用性ストレージシステムは、データ転送のための第 1 のインターフェースと、状態監視を含むシステム管理のための第 2 のインターフェースと、二重化ドライバとを有する少なくとも 1 つのホストコンピュータと、データ転送のための第 3 のインターフェースと、状態監視を含むシステム管理のための第 4 のインターフェースとを有する少なくとも 1 つのストレージ装置と、上記ホストコンピュータの第 1 のインターフェースと上記ストレージ装置の第 3 のインターフェースとを接続するデータ転送用接続インターフェースと、上記ホストコンピュータの第 2 のインターフェースと上記ストレージ装置の第 4 のインターフェースとを接続する管理用接続インターフェースとを備え、上記コンピュータの二重化ドライバが、当該コンピュータと上記ストレージ装置との間のデータ転送で、そのデータ転送用のバスとして、当該コンピュータの第 1 のインターフェースデータ転送用接続インターフェースとストレージ装置の第 3 のインターフェースの経路である第 1 のバスと、当該コンピュータの第 2 のインターフェースデータ転送用接続インターフェースとストレージ装置の第 4 のインターフェースの経路である第 2 のバスとを切り替え制御することを特徴とする。

【0009】このような構成においては、データ転送用の上記第 1 のバスに加えて、従来であれば管理データの転送のみに用いられていた管理用のバスのうちの上記第 2 のバスを、第 1 のバスと同様のデータ転送用バスとしても割り当てることができる。この結果、従来とは異なって、二重化のためのハードウェア構成、即ちホストコンピュータ及びストレージ装置内に、第 1 及び第 3 のイ

ンタフェースに相当する別のデータ転送用インターフェースを設ける必要がなく、ハードウェア構成の削減が図れ、しかも従来の構成と同等のシステムの可用性を実現できる。ここで、ホストコンピュータ及びストレージ装置内に、それぞれ上記第 1 及び第 3 のインターフェースを複数設けて、上記第 1 のバス自体を多重化するならば、より一層システムの可用性を高めることができる。

【0010】また、上記二重化ドライバによる制御のためには、ホストコンピュータ内に上記第 1 及び第 2 のバスを監視する手段を設けるとよい。

【0011】また、上記コンピュータ内に、第 2 のインターフェース及び管理用接続インターフェースを利用しての管理データ転送と、管理用接続インターフェースを含む第 2 のバスを利用してのストレージ装置との間のデータ転送とを制御する管理用ドライバを設けるとよい。同様に、上記ストレージ装置の例えばコントローラ内に、第 4 のインターフェース及び管理用接続インターフェースを利用しての管理データ転送と、管理用接続インターフェースを含む第 2 のバスを利用してのホストコンピュータとの間のデータ転送とを制御する管理用ドライバを設けるとよい。

【0012】以上の構成の高可用性ストレージシステムに係る発明は、当該システム中のホストコンピュータに係る発明としても成立する。また、このホストコンピュータに係る発明は、当該コンピュータにより実行されるドライバプログラムに係る発明、及び当該コンピュータにより実行されるドライバプログラムを記憶したコンピュータ読み取り可能な記憶媒体に係る発明、更には当該コンピュータにおける二重化バス制御方法に係る発明としても成立する。

【0013】

【発明の実施の形態】以下、本発明の実施の形態につき図面を参照して説明する。

【0014】図 1 は本発明の一実施形態に係る高可用性ストレージシステムの構成を示すブロック図である。図 1 のシステムは、主として、n 台のホスト (ホストコンピュータ) C 1 1 ~ C 1 n と、ストレージ装置としてのディスクアレイ装置 C 2 と、管理端末 C 3 とから構成される。

【0015】ホスト C 1 1 ~ C 1 n 及びディスクアレイ装置 C 2 は、SAN (Storage Area Network : ストレージエリアネットワーク) 1 0 により相互接続されている。これによりホスト C 1 1 ~ C 1 n は、SAN 1 0 を介してディスクアレイ装置 C 2 を共有することができる。

なお、図 1 では、ディスクアレイ装置 C 2 は 1 台であるが、複数のディスクアレイ装置 C 2 を SAN 1 0 を介してホスト C 1 1 ~ C 1 n と接続することも可能である。

【0016】また、ホスト C 1 1 ~ C 1 n 及びディスクアレイ装置 C 2 は、LAN (Local Area Network : ローカルエリアネットワーク) 2 0 により管理端末 C 3 と接

管理用インターフェースを用いた高可用性ストレージシステム、同システムに用いられるホ...

特開2002-229740

続されている。

【0017】ホストC1x(x=1~n)は、データ転送用インターフェースP1及び管理用インターフェースP2の両ハードウェア構成要素を有している。データ転送用インターフェースP1は、ホストC1xとディスクアレイ装置C2とをSAN10にて接続するためインターフェースを提供する。管理用インターフェースP2は、ホストC1xと管理端末C3とをLAN20にて接続するためのインターフェースを提供する。

【0018】ディスクアレイ装置C2は、複数のディスク装置(図示せず)と、これらのディスク装置のアクセス制御及びディスクアレイ装置C2全体の制御を行うコントローラ(ディスクアレイコントローラ)P3とから構成される。コントローラP3は、データ転送用インターフェースP4及び管理用インターフェースP5の両ハードウェア構成要素を有している。データ転送用インターフェースP4は、ディスクアレイ装置C2とホストC1xとをSAN10にて接続するためインターフェースを提供する。管理用インターフェースP5は、ディスクアレイ装置C2と管理端末C3とをLAN20にて接続するためのインターフェースを提供する。

【0019】管理端末C3は、ホストC11~C1n及びディスクアレイ装置C2の状態の監視を含むシステム管理を集中して実行する管理専用の端末である。管理端末C3は、当該端末C3のハードウェア構成要素としての管理用インターフェースP6を有している。管理用インターフェースP6は、管理端末C3とホストC11~C1n及びディスクアレイ装置C2とをLAN20にて接続するためのインターフェースを提供する。

【0020】ホストC11~C1nの各データ転送用インターフェースP1には、それぞれデータ転送用インターフェースケーブルF11~F1nの一端が接続され、ディスクアレイ装置C2のデータ転送用インターフェースP4には、データ転送用インターフェースケーブルF2の一端が接続されている。ケーブルF11~F1n、F2はファイバチャネルケーブル等である。SAN10は、ケーブルF11~F1n、F2の他端をデータ転送用接続インターフェースとしてのファブリックスイッチに接続することにより実現される。つまり、SAN10は実質的にデータ転送用接続インターフェースである。

【0021】ホストC11~C1nの管理用インターフェースP2には、それぞれ管理用インターフェースケーブルE11~E1nの一端が接続され、ディスクアレイ装置C2の管理用インターフェースP5には、管理用インターフェースケーブルE2の一端が接続されている。また、管理端末C3の管理用インターフェースP6には、管理用インターフェースケーブルE3の一端が接続されている。ケーブルE11~E1n、E2、E3はCSMA/CD方式を適用するLANケーブルである。LAN20は、ケーブルE11~E1n、E2、E3の他端を管理用接続

インターフェースとしてのハブに接続することにより実現される。つまり、LAN20は実質的にデータ転送用接続インターフェースである。

【0022】図2は、ホストC1x(x=1~n)の主としてソフトウェア構成を示す。ホストC1xには、アプリケーションソフトウェアS1、オペレーティングシステム(OS)S2、二重化ドライバS3、SCSIドライバS4、S41、データ転送用ドライバS5、及び管理用ドライバS6の各ソフトウェアがインストールされている。ドライバS3~S6からなるドライバ部(ドライバソフトウェア)S7は、例えばCD-ROM等の記憶媒体からホストC1x内の記憶装置に読み込まれたものである。このドライバ部S7が、インターネット等を介してホストC1x内の記憶装置にダウンロードされたものであっても構わない。

【0023】二重化ドライバS3は、SCSIドライバS4、データ転送用ドライバS5及びデータ転送用インターフェースP1から、図1中のケーブルF1x及びF2を経由してディスクアレイ装置C2のデータ転送用インターフェースP4までのパス(以下、パスFと称する)と、SCSIドライバS41、管理用ドライバS6及び管理用インターフェースP2から、ケーブルE1x及びE2を経由してディスクアレイ装置C2の管理用インターフェースP5までのパス(以下、パスEと称する)の切り替えを制御する。このように本実施形態の特徴は、状態監視を含む管理用のパスEが、ホストC1xとディスクアレイ装置C2との間のデータ転送用にも切り替えられる点にある。

【0024】SCSIドライバS4は、SCSIデバイスとして認識されるディスクアレイ装置C2とホストC1xとの間のデータ転送に関する、オペレーティングシステムS2とデータ転送用ドライバS5との間の制御を行う。同様に、SCSIドライバS41は、ディスクアレイ装置C2とホストC1xとの間のデータ転送に関する、オペレーティングシステムS2と管理用ドライバS6との間の制御を行う。なお、データ転送用ドライバS5及び管理用ドライバS6に、それぞれSCSIドライバS4及びS41の機能を持たせることも可能である。

【0025】データ転送用ドライバS5は、ホストC1xのデータ転送用インターフェースP1を制御し、SCSIドライバS4にSCSIデバイスとしてディスクアレイ装置C2を認識させるための制御及びデータの授受を行う。管理用ドライバS6は、ホストC1xの管理用インターフェースP2を制御し、SCSIドライバS41にSCSIデバイスとしてディスクアレイ装置C2を認識させるための制御及びデータの授受を行う。また管理用ドライバS6は、オペレーティングシステムS2に対して、通常のネットワーク機器のインターフェースとして認識させるための制御を行う。

【0026】図3は、ディスクアレイ装置C2内の主と

してコントローラP3のソフトウェア構成（例えばファームウェア構成）を示す。コントローラP3には、オペレーティングシステムS8、システム制御S9、アクセス制御S10、状態監視S11、データ転送用ドライバS12及び管理用ドライバS13の各ソフトウェアがインストールされている。

【0027】オペレーティングシステムS8はディスクアレイ装置C2のコントローラP3に組み込まれた基本ソフトウェアであり、例えばROM等の不揮発性メモリに予め格納されている。

【0028】システム制御S9は、ディスクアレイ装置C2全体の制御を行う。アクセス制御S10は、ディスク装置へのアクセスを制御する。状態監視S11は、ディスクアレイ装置C2の電源、ファン、ディスク装置に代表される構成要素の状態、及び運用状況など、ディスクアレイ装置C2の種々の状態を監視する。

【0029】データ転送用ドライバS12は、ディスクアレイ装置C2のデータ転送用インターフェースP4とアクセス制御S10との間のデータの授受を制御する。管理用ドライバS13は、ディスクアレイ装置C2の管理用インターフェースP5を制御し、状態監視のときは、当該インターフェースP5と状態監視S10との間のデータの授受を制御し、ディスクアクセスのときは、当該インターフェースP5とアクセス制御S9との間のデータの授受を制御する。

【0030】次に、本実施形態の動作を、（1）ホストの動作と（2）ディスクアレイ装置の動作に分けて順次説明する。

【0031】（1）ホストの動作

まず、ホストC1x（x=1～n）に図2中の二重化ドライバS3が実装されていないものとすると、そのホストC1xのオペレーティングシステムS2からは、ディスクアレイ装置C2がバスFとバスEの2つのバスによって2台のディスクアレイ装置として認識される。

【0032】これに対して本実施形態のように二重化ドライバS3が実装されたホストC1xでは、その二重化ドライバS3により以下の制御が行われる。

【0033】まず二重化ドライバS3は、バスF及びバスEの各々の経路から認識されるディスクアレイ装置が、同じディスクアレイ装置であることを確認するための処理を行う。この確認は、バスF及びバスEをそれぞれ介してディスクアレイ装置から固有の情報を読み出すことで行う。ディスクアレイ装置固有の情報とは、ディスクアレイ装置に付けられたシリアルID番号、ファイバチャネル規格で定められたファイバチャネル機器のWWN（World Wide Name）など、ディスクアレイ装置を特定することが可能な情報である。なお、SAN10を介して複数のディスクアレイ装置C2が接続されている場合には、各ディスクアレイ装置C2毎に固有の情報を読み出せばよい。

【0034】次に二重化ドライバS3は、バスF及びバスEからそれぞれ認識される同じディスクアレイ装置C2を、仮想の1台のディスクアレイ装置（仮想デバイス）Dxとして、オペレーティングシステムS2に認識させる。

【0035】そして、オペレーティングシステムS2から仮想デバイスDxに対するディスクアクセスの要求があった場合、二重化ドライバS3は、バスFが正常に動作しているならば、当該バスFを用いてディスクアレイ装置C2へのアクセスを行う。これに対し、ケーブルF1xの断線などの要因でバスFに何らかの障害が発生したことの通知を、データ転送用インターフェースP1、データ転送用ドライバS5及びSCSIドライバS4を経由して受けたときは、二重化ドライバS3はオペレーティングシステムS2に障害の発生を通知し、バスEを用いてディスクアレイ装置C2へのアクセスを行う。

【0036】二重化ドライバS3は、当該ドライバS3の設定情報に設定された時間間隔で定期的にバスF及びバスEの障害監視を行い、バスの異常を検出した場合は、オペレーティングシステムS2に通知を行い、障害の発生していないバスを用いてディスクアクセスを行う。

【0037】また二重化ドライバS3は、バスの障害が回復されたことを検出した場合には、その旨をオペレーティングシステムS2に通知する。もし、バスFで発生していた障害が回復した場合には、オペレーティングシステムS2はバスEの経路を介して行っていたディスクアレイ装置C2へのディスクアクセスを、バスFの経路でのディスクアクセスに切り替える。

【0038】ホストC1xの管理用ドライバS6は、管理用インターフェースP2を制御し、オペレーティングシステムS2から制御されるホストC1xの状態を、ケーブルE11及びE3を経由して管理端末C3に通知する。この他に管理用ドライバS6は、本発明に関係する以下の制御を行う。

【0039】即ちホストC1xの管理用ドライバS6は、当該ホストC1xの二重化ドライバS3、及びSCSIドライバS41から制御されるSCSIプロトコルのディスクアクセスコマンド（例えばリード／ライトコマンド）、ディスクアクセスデータを、LAN20を介して転送するための、i SCSI、SCSI over IPなどと呼ばれるデータ形式（i SCSI規格）に変換し、管理用インターフェースP2を制御してLAN20への転送を行う。また管理用ドライバS6は、LAN20から管理用インターフェースP2が受信した、ディスクアレイ装置C2の管理用インターフェースP5から転送されたi SCSIのデータをデータ変換して、SCSIドライバS41に転送する。

【0040】（2）ディスクアレイ装置の動作

ディスクアレイ装置C2内の管理用ドライバS13は、

管理用インターフェースP5を制御し、状態監視S11から制御されるディスクアレイ装置C2の状態をケーブルE2及びE3を経由して管理端末C3に通知する。この他に管理用ドライバS13は、本発明に関係する以下の制御を行う。

【0041】即ちディスクアレイ装置C2の管理用ドライバS13は、当該ディスクアレイ装置C2のアクセス制御S10から制御されるSCSIプロトコルのディスクアクセスコマンド（例えば応答コマンド）、ディスクアクセスデータを、LAN20を介して転送するためのデータ形式（iSCSI規格）に変換し、管理用インターフェースP5を制御してLAN20への転送を行う。また管理用ドライバS13は、LAN20から管理用インターフェースP5が受信した、ホストC1xの管理用インターフェースP2から転送されたiSCSIのデータをデータ変換して、アクセス制御S10に転送する。

【0042】次に、上記したホストC1xの二重化ドライバS3の動作の詳細を、(a)バスが二重化されたデバイスの認識処理、(b)ディスクアクセス処理、(c)バスの障害監視処理の各々に分けて説明する。

【0043】(a)バスが二重化されたデバイスの認識処理

まず、ホストC1xの二重化ドライバS3による、バスが二重化されたデバイス（ディスクアレイ装置C2）の認識処理、即ちバスF及びバスEの各々の経路から認識されるディスクアレイ装置が、同じディスクアレイ装置であることを確認するための処理について、図4のフローチャートを参照して説明する。

【0044】今、二重化ドライバS3が、アイドル状態（ステップ100）において、オペレーティングシステムS2からのディスク接続情報の要求を受信したものとする（ステップ101）。

【0045】二重化ドライバS3は、ディスク接続情報要求に応じて、バスFに接続されたディスクアレイ装置の固有の情報（ディスク識別情報）、及びバスEに接続されたディスクアレイ装置の固有の情報を、それぞれバスF及びバスEの経路から読み出す。

【0046】次に二重化ドライバS3は、バスFから読み出したディスク装置固有の情報と、バスEから読み出したディスク装置固有の情報を比較して、両情報が一致しているか否かにより、バスFとバスEとが同じディスクアレイ装置に接続されているか否かを判定する（ステップ103）。

【0047】もし、バスFとバスEとが同じディスクアレイ装置に接続されていないと判定した場合、即ちバスFとバスEとが異なるディスクアレイ装置に接続されていると判定した場合、二重化ドライバS3はバスの二重化がされていないものとして、ステップ104、105を実行する。二重化ドライバS3はこのステップ104、105において、バスFのディスクアレイ装置をD

x1とし、バスEのディスクアレイ装置をDx2とする。

【0048】これに対し、バスFとバスEとが同じディスクアレイ装置に接続されていると判定した場合には、

05 即ちバスFに接続されたディスクアレイ装置とバスEに接続されたディスクアレイ装置とが同一のディスクアレイ装置であると判定した場合には、二重化ドライバS3は当該ディスクアレイ装置との間でバスの二重化がされているものとして、ステップ106を実行する。二重化
10 ドライバS3はこのステップS106において、バスFに接続され、且つバスEにも接続された同一のディスクアレイ装置（ここではディスクアレイ装置C11）を仮想のディスクアレイ装置（仮想デバイス）Dx1とする。

15 【0049】そして二重化ドライバS3は、認識したデバイス（ディスクアレイ装置）の情報Dxm（m=1またはm=1, 2）をオペレーティングシステムS2に返してアイドル状態（ステップ100）に戻る。

【0050】(b)ディスクアクセス処理

20 次に、二重化ドライバS3によるディスクアクセス処理について、図5のフローチャート（のステップ201～209）を参照して説明する。

【0051】まず二重化ドライバS3が、アイドル状態（ステップ100）において、オペレーティングシステムS2からのDx1（で示されるディスクアレイ装置）に対するディスクアクセス要求を受信したものとする（ステップ201）。

【0052】二重化ドライバS3は、Dx1への実行バス（バスFとする）の状態が正常であるか否かを判定する（ステップ202）。この判定は、後述する障害監視処理（ステップ301～306）での前回の実行結果として、エラーが通知されているか否かに基づいて行われる。

【0053】二重化ドライバS3は、ステップ202で実行バスの状態が異常であると判定した場合、Dx1へのバスが二重化されているか否かを、先のデバイス認識処理（ステップ101～107）でのDxyの設定内容から判定する（ステップ203）。

【0054】もし、Dx1へのバスが二重化されているならば、二重化ドライバS3は待機バス（バスEとする）の状態が正常であるか否かを判定する（ステップ204）。そして二重化ドライバS3は、待機バスが正常であるならば、待機バスと実行バスとを切り替えて、今まで待機バスであったバス（バスE）を実行バスに、実行バスであったバス（バスF）を待機バスに変更する（ステップ205）。

【0055】これに対し、ステップ204で待機バスも異常であると判定した場合には、ディスクアクセス不可能であることから、二重化ドライバS3はアクセスが実行できなかったことを示すアクセス異常完了をオペレ

ティングシステムS2に通知する（ステップ209）。また、ステップ203でDx1へのバスが二重化されていないと判定した場合にも、ディスクアクセス不可能であることから、二重化ドライバS3はステップ209を実行して、アクセス異常完了をオペレーティングシステムS2に通知する。

【0056】さて、ステップ202で実行バスが正常であると判定した場合、或いはステップ205で待機バスと実行バスとを切り替えた場合、二重化ドライバS3は現在の実行バスを介してオペレーティングシステムS2から要求されたディスクアクセスを実行する（ステップ206）。

【0057】二重化ドライバS3は、ステップ206でのディスクアクセスが正常に実行された場合（ステップ207のYES）、正常にアクセスが完了したことをオペレーティングシステムS2に通知して（ステップ208）、アイドル状態（ステップ100）に戻る。これに対し、ステップ206でのディスクアクセスの実行中に実行バスの異常を検出した場合やその他の要因でディスクアクセスが正常に実行されなかった場合には（ステップ207のNO）、二重化ドライバS3は待機バスでのディスクアクセスのためにステップ203に分岐する。

【0058】(c) パスの障害監視処理

次に、二重化ドライバS3によるバスの障害監視処理について、図5のフローチャート（のステップ301～306）を参照して説明する。

【0059】二重化ドライバS3は、アイドル状態（ステップ100）において一定時間が経過する毎に、バスの障害を監視する処理（ステップ301～306）を実行する。即ち二重化ドライバS3は、バスの正常／異常を定期的に監視する。この監視は、二重化ドライバS3の設定情報に設定された時間間隔で行われる。

【0060】まず二重化ドライバS3は、バスの障害監視を実行する（ステップ301）。この障害監視、即ち実行バスが正常か否かの監視は、Dx1への実行バスにコマンドを送信することで行われる。このコマンドには、ディスクデータに影響を及ぼす恐れのないステータス読み出しなどのコマンド、例えばSCSIプロトコルのMode Senseコマンドを用いる。

【0061】二重化ドライバS3は、実行したコマンドが正常に終了したか否かにより、Dx1への実行バスが正常であるか否かを判定する（ステップ302）。もし、実行したコマンドが正常終了した場合、二重化ドライバS3は実行バスが正常であると判定して処理を完了し、アイドル状態（ステップ100）に戻る。

【0062】これに対し、実行したコマンドが異常終了した場合には、二重化ドライバS3はDx1への実行バスが異常であると判定する。この場合、二重化ドライバS3は前記ステップ203と同様に、Dx1へのバスが二重化されているか否かを判定する（ステップ303）

もし、Dx1へのバスが二重化されている場合には、二重化ドライバS3は前記ステップ204と同様に、待機バス（バスEとする）の状態が正常であるか否かを判定する（ステップ304）。そして待機バスが正常であるならば、二重化ドライバS3は前記ステップ205と同様に、待機バスと実行バスとを切り替えて、今まで待機バスであったバスを実行バスに、実行バスであったバスを待機バスに変更する（ステップ305）。

【0063】一方、Dx1へのバスが二重化されていない場合、或いは二重化されていても待機バスが異常な場合には、二重化ドライバS3は障害監視処理にてバスの障害を検出したことをオペレーティングシステムS2に通知して（ステップ306）、アイドル状態（ステップ100）に戻る。

【0064】以上に述べたように本実施形態においては、ホスト（ホストコンピュータ）C1xの二重化ドライバS3に、従来のSCSIドライバS4経由でのデータ転送用ドライバS5の制御に加えて、SCSIドライバS41経由での管理用ドライバS6の制御を行う機能（以下、機能1と称する）を付加したことを特徴とする。

【0065】次に、ホストC1xの管理用ドライバS6に、従来のLAN20を利用しての管理データ（監視データ）転送に加えて、LAN20を介してSCSIプロトコルのデータを転送する機能（以下、機能2と称する）を付加したことをも特徴とする。

【0066】そして、ディスクアレイ装置C2の管理用ドライバS13に、従来のLAN20を利用しての管理データ（監視データ）転送に加えて、LAN20を介してSCSIプロトコルのデータを転送する機能（以下、機能3と称する）を付加したことをも特徴とする。

【0067】このような機能1乃至3の追加により、データ転送用のバスFに加えて、従来であれば管理データの転送のみに用いられていた管理用のバスのうち、ホストC1xの管理用インターフェースP2からケーブルE1x、LAN20及びケーブルE2を介してディスクアレイ装置C2の管理用インターフェースP5に至るバスEを、バスFと同様のデータ転送用バスとしても割り当てることができる。この結果、従来の高可用性ストレージ

システムの二重化構成とは異なって、二重化のためのハードウェア構成、即ちホストC1x及びディスクアレイ装置C2内に、データ転送用インターフェースP1及びP4に相当するもう1つのデータ転送用インターフェースを設けることも、データ転送用インターフェースケーブルF1x及びF2に相当するデータ転送用インターフェースケーブルを設けることも必要なく、ハードウェア構成の削減が図れ、しかも従来の構成と同等のシステムの可用性を実現できる。

【0068】また、従来の二重化構成を適用したシステムに、上記機能1乃至3を追加して管理用バスをデータ

転送用バスにも利用できるようにした場合には、バスの三重化構成が実現できるため、同じハードウェア構成でありながら、より可用性の高いシステムを実現することができる。

【0069】以上の実施形態では、独立の管理端末C3が設けられている場合について説明したが、これに限るものではない。例えば、ホストC11～C1nのいずれか1つに、ホストC11～C1n及びディスクアレイ装置C2の状態監視を含むシステム管理を行う機構を持たせるようにしてもよい。本発明では、このようなホスト（管理用ホスト）内に設けられるシステム管理機構も、管理端末と定義する。

【0070】なお、本発明は、上記実施形態に限定されるものではなく、実施段階ではその要旨を逸脱しない範囲で種々に変形することが可能である。更に、上記実施形態には種々の段階の発明が含まれており、開示される複数の構成要件における適宜な組み合わせにより種々の発明が抽出され得る。例えば、実施形態に示される全構成要件から幾つかの構成要件が削除されても、発明が解決しようとする課題の欄で述べた課題が解決でき、発明の効果の欄で述べられている効果が得られる場合には、この構成要件が削除された構成が発明として抽出され得る。

【0071】

【発明の効果】以上詳述したように本発明によれば、データ転送用のバス（第2のバス）に加えて、従来であれば管理データの転送のみに用いられていた管理用接続インターフェースを含む管理用のバスのうちの、ホストコンピュータの管理用インターフェース（第2のインターフェース）から当該管理用接続インターフェースを経由してストレージ装置の管理用インターフェース（第4のインターフェース）に至るバス（第1のバス）をデータ転送用バスとしても割り当てることができる。この結果、データ転送用バスを二重化するのに、二重化のためのハードウェア構成を必要とせず、したがってハードウェア構成の削減が図れ、しかも従来の構成と同等のシステムの可用性を

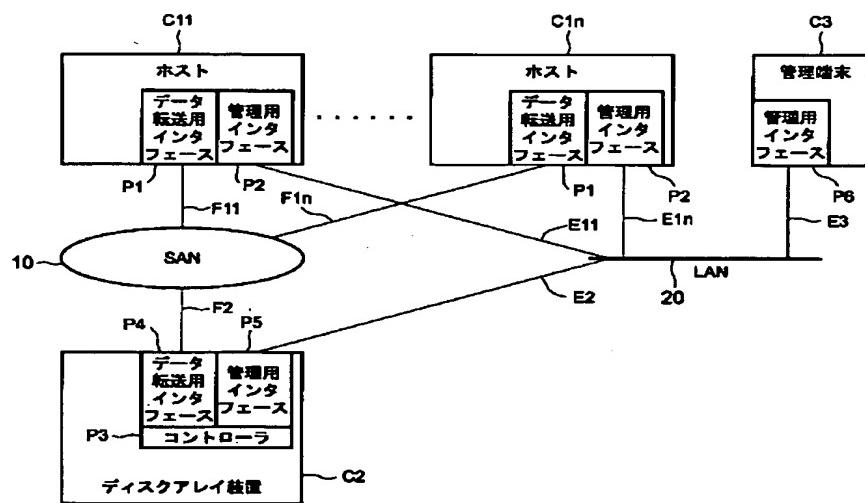
実現できる。

【図面の簡単な説明】

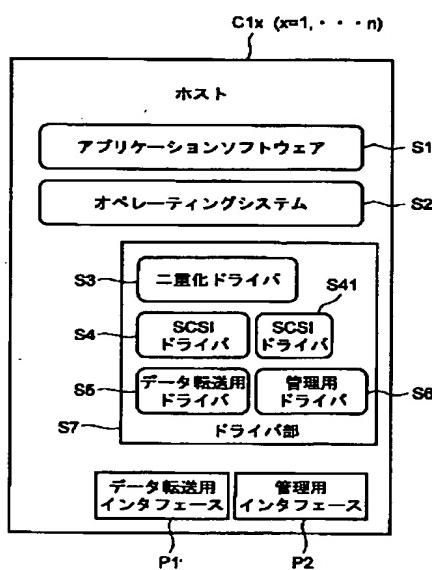
【図1】本発明の一実施形態に係る高可用性ストレージシステムの構成を示すブロック図。

- 05 【図2】図1中のホストC1x（x=1～n）の主としてソフトウェア構成を示す図。
【図3】図1中のディスクアレイ装置C2内の主としてコントローラP3のソフトウェア構成を示す図。
【図4】図3中のホストC1xの二重化ドライバS3による、バスが二重化されたデバイスの認識処理を説明するためのフローチャート。
【図5】図3中のホストC1xの二重化ドライバS3による、ディスクアクセス処理及びバスの障害監視処理を説明するためのフローチャート。
- 15 【符号の説明】
C11～C1n, C1x…ホスト（ホストコンピュータ）
C2…ディスクアレイ装置（ストレージ装置）
C3…管理端末
20 P1…データ転送用インターフェース（第1のインターフェース）
P2…管理用インターフェース（第2のインターフェース）
P3…コントローラ
P4…データ転送用インターフェース（第3のインターフェース）
25 P5…管理用インターフェース（第4のインターフェース）
P6…管理用インターフェース
S2, S8…オペレーティングシステム
S3…二重化ドライバ（監視手段、切り替え制御手段）
30 S5, S12…データ転送用ドライバ
S6…管理用ドライバ（転送制御手段）
S13…管理用ドライバ
10…SAN（ストレージエリアネットワーク、データ転送用接続インターフェース）
35 20…LAN（ローカルエリアネットワーク、管理用接続インターフェース）

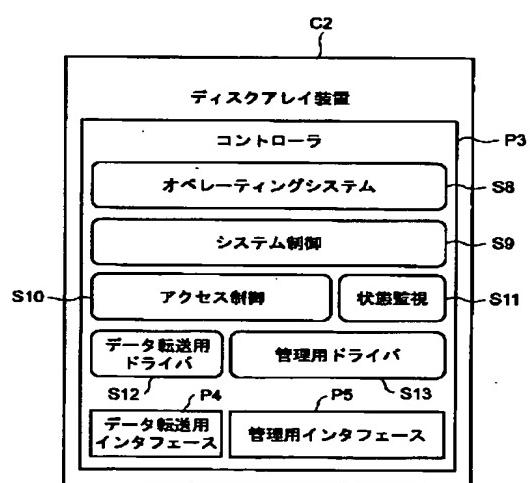
【図1】



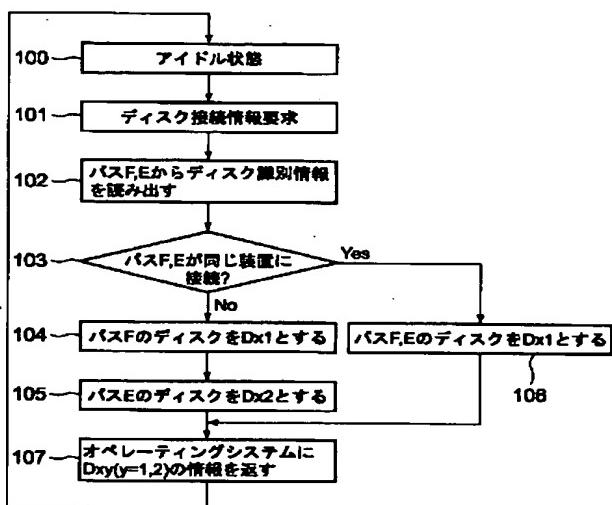
【図2】



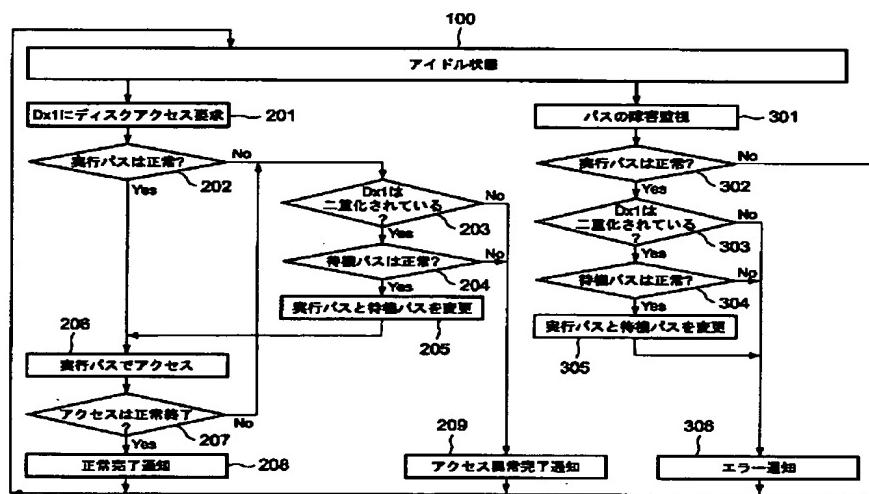
【図3】



【図4】



【図5】



フロントページの続き

- (54) 【発明の名称】 管理用インターフェースを用いた高可用性ストレージシステム、同システムに用いられるホストコンピュータ、及び同コンピュータにより実行されるドライバプログラムを記憶したコンピュータ読み取り可能な記憶媒体